

DOCUMENT RESUME

ED 314 500

TM 014 425

AUTHOR Crehan, Kevin D.; And Others
TITLE An Investigation of the Optimal Number of Options and the Inclusive "None of These" Option.
PUB DATE 89
NOTE 10p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *College Students; Higher Education; *Item Analysis; *Multiple Choice Tests; Psychology; Test Construction; *Test Format; Test Items; Test Reliability

ABSTRACT

Two issues in the writing of multiple-choice test items were investigated: a comparison of three versus four options; and the use of the inclusive "none of these" option versus a content option. Subjects were 270 introductory psychology students, who were enrolled at a large southwestern university, responding to a final examination in psychology composed of 48 items. Each item was written in four versions: (1) four options without "none of these"; (2) four options with "none of these"; (3) three options without "neither of these"; and (4) three options with "neither of these". Item analysis and test analysis comparing the manipulated item versions were conducted using item response theory. The three-option format was found to be less difficult than the four-option format, and the use of the "none of these" option resulted in more difficult items. There was no difference in discrimination between three- and four-option formats, a possible argument in favor of a three-option format. However, the study did indicate observed differences in the reliability favoring the four-option format. (SLD)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

ED314500

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☒ This document has been reproduced as
received from the person or organization
originating it

☐ Minor changes have been made to improve
reproduction quality

• Points of view or opinions stated in this docu-
ment do not necessarily represent official
OERI position or policy

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

KEVIN D. CREHAN

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

An Investigation of the Optimal Number
of Options and the Inclusive
"None of These" Option

Kevin D. Crehan
University of Nevada, Las Vegas

Thomas M. Haladyna
Arizona State University West Campus

Britton Brewer
Arizona State University

ABSTRACT

Two issues in the writing of multiple-choice test item are investigated in this study: (1) a comparison of three- verses four-options and, (2) use of the inclusive "none of these" option verses a content option. Subjects are two hundred twenty introductory psychology students responding to a final exam composed of forty-eight items. Each item is written in four versions: (1) four options without "none of these", (2) four-options with "none of these", (3) three options without "neither of these", and (4) three options with "neither of these". Item and test analysis comparing the manipulated item versions will be conducted using methods of item response theory.

TM014425

An Investigation of the Optimal Number
of Options and the Inclusive
"None of These" Option

Kevin D. Crehan
University of Nevada, Las Vegas

Thomas M. Haladyna
Arizona State University West Campus

Britton Brewer
Arizona State University

Item writing is more art than science. The technology for multiple-choice test item writing is relative stagnant despite rapid advances in statistical theories of test scores.

This study examined two major issues in the design of test items: the desirable number of options for a multiple-choice test item and use of the inclusive 'none of these' option.

Related Research

Number of Options

A comprehensive review of the literature on the optimal number of items by Haladyna and Downing (1989a) revealed that generally more options are better, but an empirical study by Haladyna and Downing (1989b) revealed the average number of useful distractors in a well-developed test of medical knowledge was two! Is the item writer and the resultant test at an advantage for developing more than a three-option item? Or can we expect only two useful distractors per on the average. Owen and Froman (1987) convincingly showed that information yield, difficulty, and discrimination were not different for comparable three-option and five-option tests, yet their review of textbooks (similar to the study conducted by Haladyna and Downing (1989a), revealed that most textbook authors recommend four or five options. Thus evidence is accumulating to suggest that designers and item writers of standardized tests and teacher-made tests need only write three options (the right answer and the two distractors) for the reasons provided in this review of literature.

None of These

In a review of 46 references dealing with multiple-choice item writing, Haladyna & Downing (1989a) found that 34 (73%) of these references stated support or lack of support for the rule "Avoid the use of the 'None of these' option. This was the tenth most often mentioned rule, and this survey was taken as evidence of the importance of the rule for item writers. However, authors

were divided on their support for this rule, with 19 for and 15 against. Obviously some controversy exists in the validity of the rule.

Empirical research on this item writing rule has been limited to only ten studies (Boynton, 1950; Dudycha & Carpenter, 1973; Forsyth & Sprastt, 1980; Hughes & Trimble, 1965; Mueller, 1975; Oosterhof & Coats, 1984; Rimland, 1960; Schmeiser & Whitney, 1975; Wesman & Bennett, 1946; Williamson & Hopkins, 1967). All of these studies involved the item characteristic of difficulty, but only five studied item discrimination and reliability, and only two validity. In all instances, the use of none of the above option made items more difficult, the mean effect across nine studies where results were aggregable was 4.8%. With discrimination, avoiding the inclusive "none of these" option made items slightly more discriminating, .03, while reliability was improved by a factor of .04.

The present study was designed to gather additional evidence concerning the effectiveness of the three-option test item as compared to the four-option test item and the inclusive "none of these" option as compared to a content option.

Method

Subjects. Two hundred-twenty introductory psychology students at a large southwestern university responded to the instruments used in this study.

Instruments. Forty-eight items from the instructor's manual for Myer's (1986) text entitled Psychology were selected by the course instructor and were rewritten in four versions:

1. four options without "none of these" option (S4),
 2. four options with "none of these" option (I4),
 3. three options without "neither of these" option (S3),
- and
4. three options with "neither of these" option (I3).

An example of the manipulations for one item is presented in Figure 1.

Insert Figure 1 about here

When reducing the number of options for an item, the least functional distractors were eliminated, i.e., distractors with the lowest discrimination indices based on available item data. Items were key balanced and four forms of the final course exam were developed with each form containing twelve items written in each of the four versions. Therefore, each form of the test was composed of forty-eight items measuring the same content specifics but with item format manipulated as indicated.

Data collection. Test forms were administered to two sections of an introductory psychology course as the final exam with approximately fifty-five examinees responding to each form.

Figure 1

The following is an example of an item written in each of the four formats employed in this study.

- S4 - A therapist who uses a mixture of different techniques is considered to be:
- a. a cognitive-behavior therapist.
 - b. a transactional analyst.
 - *c. eclectic.
 - d. client-centered.
- I4 - A therapist who uses a mixture of different techniques is considered to be:
- a. a cognitive-behavior therapist.
 - b. a transactional analyst.
 - *c. eclectic.
 - d. none of these
- S3 - A therapist who uses a mixture of different techniques is considered to be:
- a. a cognitive-behavior therapist.
 - b. a transactional analyst.
 - *c. eclectic.
- I3 - A therapist who uses a mixture of different techniques is considered to be:
- a. a cognitive-behavior therapist.
 - b. a transactional analyst.
 - *c. neither of these

Results and Analysis

Mean item difficulties and discriminations for the forty-eight items under each item format are reported in table 1. Table 1 also presents the reliabilities of the twelve item subscales averaged over test forms.

Insert table 1 about here

Repeated measures contrasts comparing mean difficulties of the three option to the four option item format and the "none of these" to the specific option format were tested at the .05 level. The mean difficulties for the 24 items under three and four option formats were .772 and .741 respectively. The difference of .031 was significant ($t=2.34$, $p=.024$, .95 C.I.:.004,.058). Mean difficulties for the specific option and "none of these" item formats were .784 and .728 respectively. The difference of .056 was significant ($t=3.8$, $p<.001$, .95 C.I.:.026,.085).

The same analysis was conducted for item discriminations. The mean discriminations for the three and four option formats were .339 and .363 respectively. The difference of -.024 was not significant ($t=1.2$, $p=.229$). Mean discriminations for the specific option and "none of these" item formats were .354 and .347 respectively. The difference of .007 was not significant ($t=.326$, $p=.746$).

Discussion

It was found that the three option item format was less difficult than the four option item format and that use of the "none of these" option resulted in more difficult items. These results are neither particularly surprising nor interesting. Of greater interest and importance were the comparisons of item discriminations. It was expected that use of the "none of these" option would result in lower discrimination indices consistent with some previous research. The absence of a difference in discrimination between the three and four option item formats may be interpreted as an argument favoring the three option format.

There are advantages in the three-option test item. First and foremost, item writing is less laborious. Since classroom testing and standardized testing is increasing with the need and emphasis to align teaching and testing, items always need to be developed, and the effort we expend should be minimal yet adequate. Second, the time needed to administer three-option items is less than with four-option and five-option items (Budescu & Nevo, 1987; Owen & Froman, 1987). Additionally, a recent review (Haladyna and Downing, 1989b) has shown that item writers typically do not develop effective third and fourth distractors anyway. Finally, if there is no serious loss in discrimination and test score reliability as a result of using

TABLE 1

Mean item difficulties and discriminations (standard deviations in parentheses) for forty-eight items under each item format and reliabilities of the twelve item subscales averaged over test forms.

	S4	I4	S3	I3
Difficulty	.766 (.16)	.716 (.17)	.803 (.13)	.741 (.15)
Discrimination	.357 (.16)	.368 (.14)	.352 (.18)	.522 (.13)
Reliabilities	.61	.60	.58	.52

three versus four or five options, the advantages of the three option format seem clear. Unfortunately, the present study, while failing to show differences in item discrimination, does show observed differences in reliability favoring the four option format. Median reliabilities for the eight subscales using each format were .617 versus .564 favoring the four option format.

Additional analysis will be conducted using methods of item response theory in an attempt to obtain more definitive answers to the research questions.

REFERENCES

- Boynton, M. (1950). Inclusion of "none of these" makes spelling items more difficult. Educational & Psychological Measurement, 10, 431-432.
- Budescu, D. V. (1985). Optimal number of options: An investigation of the assumption of proportionality. Journal of Educational Measurement, 22, 183-196.
- Dudycha, A. L., & Carpenter, J. B. (1973). Effects of item formats on item discrimination and difficulty. Journal of Applied Psychology, 58, 116-121.
- Forsyth, R. A. & Spratt, K. F. (1980). Measuring problem solving ability in mathematics with multiple choice items. The effect of item format on selected item and text characteristics. Journal of Educational Measurement, 17, 31-43.
- Haladyna, T. M. & Downing, S. M. (1989,a). A taxonomy of multiple-choice item writing rules. Applied Measurement in Education, 1, 37-50.
- Haladyna, T. M. & Downing, S. M. (1989, b). The validity of a taxonomy of multiple-choice item writing rules. Applied Measurement in Education, 1, 51-77.
- Hughes, H. H., & Trimble, W. E. (1965). The use of complex alternatives in multiple-choice items. Educational and Psychological Measurement, 25, 117-126.
- Mueller, D. J. (1975). An assessment of the effectiveness of complex alternatives in multiple choice achievement test items. Educational and Psychological Measurement, 35, 135-141.
- Myers, D. (1986). Psychology. New York, N.Y.: Worth Publishers.
- Oosterhof, A. C. & Coats, P. K. (1984). Comparison of difficulties and reliability of quantitative word problems in completion and multiple-choice item formats. Applied Psychological Measurement, 8, 287-294.
- Owen, S. V. & Froman, R. D. (1987). What's wrong with three option multiple-choice items? Educational and Psychological Measurement, 47, 513-521.
- Rimland, B. (1960). The effects of varying time limits and of using "right answer not given" in experimental forms of the U. S. Navy Arithmetic Test. Educational and Psychological Measurement, 20, 533-539.

Schmeiser, C. B., & Whitney, D. R. (1975). The effect of incomplete stems and "none of the above" foils on test and item characteristics. Paper presented at the annual meeting of the National Council on Measurement in Education, Washington, D.C.

Wesman, A. G., & Bennett, G. K. (1946). The use of 'none of these' as an option in test construction. Journal of educational Psychology, 37, 541-549.

Williamson, M. L., & Hopkins, K. D. (1967). The use of "none-of-these" versus homogeneous alternatives on multiple-choice tests: Experimental reliability and validity comparisons. Journal of Educational Measurement, 4, 53-58.